Authors:

Catherine Moroney, Larry DiGirolamo, Guangyu Zhao

Quality Designator:

- Stage 3 Validated: Feature-Referenced Radiometric Camera-by-Camera Cloud Mask (RCCM), Terrain-Referenced Snow-Ice Mask
- Stage 2 Validated: Angular Signature Cloud Mask (ASCM), Cloud Fractions
- Provisional: Resolution Corrected Cloud Fractions, Support Vector Machine (SVM) Scene Classifiers
- Beta: SVM Cirrus Fraction

This statement applies to the MISR Level 2TC Classifiers product (F07_0012) and beyond until further improvements to the MISR software are made. Quality statements covering earlier time periods may be accessed through links at the bottom of this page.

The MISR Level 2TC Classifiers software that generated these products is believed to be functioning well except where noted below. This statement highlights major known problems with the products and changes from the previous versions.

PRODUCT MATURITY

The Classifier product contains several retrieved cloud parameters and calculated cloud properties using the MISR level-2 cloud products as well as snow/ice information. Each field in the Classifier product had undergone different validation procedures at different stages. Refer to the quality designator above for the validation status of a particular field.

Overview

This Classifiers product provides terrain-referenced snow-ice mask, feature-referenced Radiometric Camera-by-camera Cloud Mask (RCCM) and Angular Signature Cloud Mask (ASCM) at 1.1 km resolution, and different measurements of cloud fractions with and without correction for pixel resolution effects, as well as the classification of cloud fractions by altitudes at a 17.6 km spatial resolution. In this version of the product, the derivation of any fields that require the cloud top height information uses as input the cloud top height and Stereoscopically-Derived Cloud Mask (SDCM) of the TC_Cloud product instead of the TC_Stereo product used in previous versions. Hence, any limitation and accuracy of the cloud top height retrieval and SDCM in the TC_Cloud product will subsequently affect the quality of fields that depend on these inputs. See the TC_CLOUD Quality Statement for details.

There are other major modification and updates to the Classifier algorithm implemented in this version. The complete algorithm details can be found in the Level 2 Cloud Classification Algorithm Theoretical Basis Document (ATBD).

KNOWN LIMITATIONS

Most of the Classifier products rely entirely on the cloud masking products. As with all cloud masking algorithms, the MISR cloud masking algorithms may not perform perfectly over regions dominated by optically thin clouds, sub-pixel clouds, or thick aerosol layers and regions contaminated with strong sunglint. Those interpreting or evaluating the performance of the Classifier products over these regions should be mindful of these caveats.

The RCCM and SDCM are used to calculate cloud fractions in many fields of the Classifiers product and the information contained in the GRP_RCCM and TC_CLOUD quality statements is of value when using the cloud fractions.

ASCM

In contrast to the RCCM, the ASCM algorithm is deemed to be inferior over snow/ice free surfaces and superior over snow/ice covered

surfaces. Global and regional analyses demonstrated that the ASCM is most effective over the poles and least effective over ice-free land surfaces. However, many problematic regions in the ASCM exist due to variations within each threshold classification (i.e., tuning a dataset to work over one location leads to a worsening in other locations). In addition, regions with heavy dust, rapidly changing ice cover, and horizontally variable cloud thickness (cloud parallax such as cumuliform clouds) tend to reduce the accuracy of the ASCM.

The ASCM is calculated using a look-up table that is indexed by sun-view geometry and scene type. Discontinuities are sometimes visible when moving to a new angular bin or a new scene type. Global cloud distribution maps made from the ASCM show the expected climatological cloud distributions. Validation via visual inspection, field campaigns, and satellite data show that the ASCM has accuracies of better than 90% for ice-free ocean and greater than 85% for other surfaces including snow and ice cover[MJ1] . A detailed validation is underway and the results will be referenced in this statement when available.

The snow and ice mask reported at 1° resolution in the monthly TASC dataset[MJ2] [CM3]   (and reprojected to a 1.1 km grid in the TC_CLASSIFIERS product[CM4]) is conservative, in that it errs on the side of labeling a location as snow-ice covered even when there is only a small amount of snow or ice present in the month. The dataset is assembled on a monthly basis from the appropriate daily data, and if a pixel had snow or ice for 4 or more days in the month, it will be listed as being snow-ice covered for the entire month. If the TASC mistakenly identifies a snow-ice free area as being snow-ice covered, fewer clouds will be detected by the ASCM and some thin clouds will be labeled as clear. If the TASC misses snow-ice that is actually present, the ASCM will tend to label clear snow-ice covered areas as cloud. Therefore, given the conservative nature of the snow-ice mask, the ASCM will tend towards missing thin clouds in areas of intermittent snow-ice coverage.

Cloud Fractions

As described in the Level-2 Cloud Classifier ATBD, the cloud fraction is defined as the fraction of pixels detected as high or low confidence cloudy by any or a combination of the three MISR cloud masks, the RCCM, ASCM, and SDCM. Since sub pixel clouds (by definition) do not fully cover a pixel, the true cloud fractions (as would be defined by a perfect cloud detector using pixels that are near infinitesimal in size) reported in the product will be overestimated in regions populated by small clouds (e.g., trade wind cumulus regions). A full discussion on this issue is given in Zhao and DiGirolamo (2006).

The RCCM-derived cloud fractions (newly added in this version of the software) are corrected for the pixel resolution effect using the techniques described in Jones *et al*. (2012). The quality of the correction is also described in Jones *et al*. (2012). The corrected cloud fractions are stored separately from the original cloud fractions in this product.

Cloud-Top Heights

The cloud top heights are calculated by a stereoscopic method and therefore retrieve the altitude of greatest spatial contrast as viewed by multiple MISR cameras. This may differ from the heights retrieved by a lidar or an IR sensor (see Stubenrauch *et al.* 2013 for an overview). The cloud optical depth threshold at which the stereo algorithm identifies thin clouds depends, in part, on the contrast of the underlying surface. Therefore, MISR can detect thinner, single-layered clouds over ocean than land. Where heights are retrieved, RMS errors have been determined to be ~ 560 m (see the TC_CLOUD quality statement).

Because the feature-referenced cloud masks (used in the cloud fraction calculations) are only produced when a corresponding good quality height is also available, dropouts in the cloud fractions will not be evenly distributed over all cloud types. The stereo matchers have greater difficulty with high, thin cloud as well as relatively featureless regions, although TC_CLOUD performs much better for these areas than the older version of the cloud product, TC_STEREO (see below). Additionally, if the orbit is deemed to be poorly registered, wind-corrected cloud top heights are not reported which will, in turn, result in all the feature-referenced cloud masks (and *CombinedFraction* fields) being set to NoRetrieval. The percentage of orbits that fall into this category ranges from about 5% in spring to 15% in fall (for 2012 data), with the oceanic paths being disproportionately affected.

CHANGES FROM PREVIOUS VERSIONS

This version of the Classifiers product now ingests the cloud-top heights from the TC_CLOUD product instead of the TC_STEREO product, fixes some algorithm issues, moves some fields from the TC_STEREO product into TC_CLASSIFIERS, and adds a new product: the resolution-corrected cloud fractions (available at 17.6 km resolution).

New Parameters

The feature-referenced RCCM (as calculated using the An camera data) is available in both with and without-wind-correction versions. It was originally available in the TC_STEREO product but is now contained in TC_CLASSIFIERS. The snow-ice information (reprojected to a 1.1 km grid) has also been moved from TC_STEREO to TC_CLASSIFIERS, and the feature-projected snow-ice mask has been removed.

Resolution-corrected cloud fractions have been added to the TC_CLASSIFIERS product for version 12. Since the number of pixels with any type of cloud in them increases with pixel size, it is important to have a "true" cloud fraction that does not depend on pixel resolution. These cloud fractions are derived from a look-up table that maps a "feature-vector" of observable values (such as the number of 1.1 km pixels in the 17.6 km region that are on the exterior of a cloud) to the true cloud fraction. More information on this approach can be found in Jones *et al*. (2012).

Impact of using new TC_CLOUD Heights as Input

The new TC_CLOUD wind-corrected heights have much better coverage than their TC_STEREO counterparts and this serves to increase the coverage of the feature-referenced RCCM and ASCM and, also increases the number of pixels that go into the calculation of the *CombinedFractionCloud-BestEstimate* and *CombinedFractionCloudHC* fields. TC_CLOUD can pick up thinner clouds than TC_STEREO and also has much better coverage in the polar regions, so these changes also feed into the *CombinedFractions* fields.

Slightly decreased coverage in the individual feature-referenced cloud masks relative to previous versions of the TC_CLASSIFIERS product may occasionally be visible at the swath edges, due to decreased co-registration quality in the L1B2 data in these areas.

Algorithm Changes

In earlier versions of the TC_CLASSIFIERS product, the cloud fractions (those in the *CloudFractions_17.6_km* grid) used to trend down at the swath edges because they were normalized by the fixed number of 1.1 km pixels present in a 17.6 km region. Therefore, areas of the block outside the edges of the visible swath were mistakenly included in the fractions. This has been fixed so now the cloud fractions are normalized only by the number of pixels visible within the swath. The exception to this rule is the fraction of NoRetrievals, which are still normalized by the total number of pixels in each 17.6 km region, so the values for these parameters do trend up at the swath edges.

The feature-referenced ASCM now uses the individual, high-resolution, wind-corrected cloud-top heights, rather than using the median height per block. The FR_ASCM is now the first field in the *ASCMParams_1.1_km* grid, and the corresponding camera, observable, and scattering angle parameters are also feature-referenced. The only terrain-referenced fields in the product are the raw forward and aft camera masks along with the combined forward-aft mask.

The feature-referenced cloud masks are only calculated when a valid cloud-top height retrieval exists for the given pixel. Since the stereo retrieval needs an overlap of cameras to retrieve heights, this means that there will be areas near the edge of the swath that have a terrain-referenced cloud mask retrieval, but no feature-referenced one. This is especially evident in the ASCM because it is calculated from the oblique cameras which skew outward towards the swath edges compared to the A cameras used in the stereo height retrieval. The terrain-referenced mask was retained in the product because this field has essentially complete coverage and is not clipped at the swath edges.

The SVM scene classifier and cirrus detection algorithms as well as the consensus cloud fields have not been modified or updated in this version. Their quality as of the last version is still applicable.

DIFFERENCES BETWEEN FIRSTLOOK AND FINAL PROCESSING[KM5] [CM6]

The MISR processing stream requires data (the snow-ice mask and the RCCM thresholds) derived from the current month's/season's processing. In order not to hold up standard production, it has been split into two streams: *Firstlook* processing uses the snow-ice and RCCM thresholds from the appropriate month/season from the previous year, while *Final* processing uses the snow-ice and RCCM thresholds for the current month/season. Parameter differences between *Firstlook* and *Final* in the Classifiers product are most visible in the RCCM and ASCM cloud masks and the downstream cloud fractions. *Firstlook* products are labeled as "TC_CLASSIFIERS_FIRSTLOOK"; *Final* products are simply called "TC_CLASSIFIERS". *Final* products are generated at the end of each season, typically in March, June, September, and January.

The ASCM uses the snow-ice mask to decide which version of the corresponding ASCM thresholds to use, so changes in the snow-ice coverage directly affect the ASCM contents. Differences between *Firstlook* and *Final* will be most pronounced in areas where the snow-ice presence for that month is variable. The TASC dataset has a spatial resolution of 1° x 1° and is updated on a monthly basis, so even the *Final* processing does not contain a perfect labeling of snow-ice covered and snow-ice free areas. A comparison of the *Firstlook* and *Final* products indicate a difference in the cloud mask of about 2% over the entire swath, increasing to 40% in specific blocks where the contents of the snow-ice mask changed significantly from the previous year to the current one.

Although the RCCM thresholds are not used directly in Classifiers processing, the feature-referenced RCCM in the Classifiers product is simply a re-referencing of the original terrain-referenced product present in the GRP_RCCM files, so all the differences in the GRP_RCCM processing between Firstlook and Final carry over to the Classifiers product.   The percentage of results that differ due to the updated thresholds varies from 2% to 10%, depending on the fraction of land in the orbit. Updating the thresholds particularly improves thin cloud detection. There is no change in the ocean thresholds because they are static.

The (new) resolution-corrected cloud fractions are calculated directly from the GRP_RCCM files, so all differences between the Firstlook and Final version of that product will filter down into the cloud fractions. Additionally, the original cloud fractions use the ASCM and feature-referenced RCCM, so they are also affected by any changes in the snow-ice mask and RCCM thresholds.

REFERENCES

Di Girolamo, L. and R. Davies (1994) "A band-differenced angular signature technique for cirrus cloud detection." Di Girolamo, L. and Davies, R. IEEE Trans GeoSci Remote Sensing, vol 32,

Jones, A. L., L. Di Girolamo, and G. Zhao (2012), Reducing the resolution bias in cloud fraction from satellite derived clear-conservative cloud masks, J. Geophys. Res., 117, D12201, doi:10.1029/2011JD017195.

Stubenrauch, C. J., and Coauthors, 2013: Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the GEWEX Radiation Panel. *Bull. Amer. Meteor. Soc.*, **94**, 1031–1049. doi: http://dx.doi.org/10.1175/BAMS-D-12-00117.1

Zhao, G. and L. Di Girolamo, 2006: Cloud fraction errors for trade wind cumuli from EOS-Terra instruments. Geophys. Res. Lett., 33 L20802, doi:10.1029/2006GL027088.

Also see the:

Statement dated September 19, 2007 for MISR Level 2 Top-of-Atmosphere/Cloud products from June 17, 2007 to September 19, 2007.

Statement dated August 22, 2006 for MISR Level 2 Top-of-Atmosphere/Cloud products from August 21, 2006 to June 17, 2007..

Statement dated February 21, 2006 for MISR Level 2 Top-of-Atmosphere/Cloud products from February 21, 2006 to August 21, 2006.

Statement dated December 1, 2005. for MISR Level 2 Top-of-Atmosphere/Cloud products from December 1, 2005 to February 20, 2006.

Statement dated May 13, 2005. for MISR Level 2 Top-of-Atmosphere/Cloud products from May 13, 2005 to November 30, 2005.

Statement dated November 28, 2004 for MISR Level 2 Top-of-Atmosphere/Cloud products from November 28, 2004 to May 12, 2005.

Statement dated February 13, 2004 for MISR Level 2 Top-of-Atmosphere/Cloud products from November 30, 2003 to November 27, 2004.
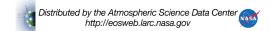
Statement dated October 20, 2003 for MISR Level 2 Top-of-Atmosphere/Cloud products from October 20, 2003 to November 30, 2003.

Statement dated August 13, 2003 for MISR Level 2 Top-of-Atmosphere/Cloud products from August 13, 2003 to October 19, 2003.

Statement dated November 12, 2002 for MISR Level 2 Top-of-Atmosphere/Cloud products from November 12, 2002 to August 12, 2003.

Statement dated April 15, 2002 for MISR Level 2 Top-of-Atmosphere/Cloud products from April 15, 2002 to November 11, 2002.

Statement dated December 3, 2001 for MISR Level 2 Top-of-Atmosphere/Cloud products from December 3, 2002 to April 14, 2002.

Statement dated September 27, 2001 for MISR Level 2 Top-of-Atmosphere/Cloud products from September 27, 2001 to December 3, 2001.

Statement dated March 30, 2001 for MISR Level 2 Top-of-Atmosphere/Cloud products from March 30, 2001 to September 27, 2001.

Statement dated February 16, 2001 for MISR Level 2 Top-of-Atmosphere/Cloud products from February 16, 2001 to March 30, 2001.

[MJ1]This sentence is repeated from the previous paragraph.

[MJ2]I was confused as the TASC is a 1° resolution, but TC_CLASSIFIERS has a 1.1 km snow-ice mask.

[CM3]Classifiers reads the 1 degree TASC data and reprojects it to a 1.1km SOM grid.

[CM4]The reprojection of the snow-ice mask to a 1.1km grid is done for ease of programming: it's nice to have a snow-ice mask on the same grid as the cloud mask when calculating the latter.

[KM5]Add sentence explaining why there are two versions (Firstlook and Final) of Classifiers.

[CM6]See paragraph below: the snow-ice masks and the RCCM thresholds.